THE USE OF ELECTRICAL COMPUTING MACHINES IN AGRICULTURAL RESEARCH 1/

1955

Walter R. Harvey 2/
Agricultural Research Service
United States Department of Agriculture

The basic electrical computing machines, subsequently referred to as ECM, were primarily developed for use in the accounting field. However, it is only within the past few years that punch card equipment has been used very extensively to summarize and analyze agricultural research data. The increasing use of ECM as a tool in research has been accompanied by the development of short cuts and wider applications in the use of these machines for this purpose. The result has generally been a large saving of time and expense to the investigator and the production of new computing machines to better do the work required.

The Punch Card

In order for the machines to perform the many functions required of them in a record keeping system or in a statistical summarization or analysis, a record of standard size and arrangement is required. The standard punch card (Figure 1) is this medium, with the holes punched in the card for each record representing the information to be processed. Once punched and checked (verified), the card (a unit record) becomes a permanent record and can be automatically processed through ECM equipment to produce the required tabulations.

The more common punch card in use has 80 vertical columns, and 12 rows. The rows consist of the digits zero through nine and two of the rows, called the 11- and 12-rows, are for "control" punching. Although there are 12 x 80 or 960 holes that can be punched in such a card, normally only one hole is punched in any column to represent a particular digit, zero through nine, and two holes are punched in any one column to represent a particular letter, a through z. In addition to being used in the alphabetical code, the 11- and 12-punch may be used with numerical data to indicate unusual facts, such as a negative number, no information available on that particular item for this card, special decks of cards, etc.

1/ The photographs in this article are intended only to illustrate the various types of cards and electrical computing equipment discussed. Their use does not imply endorsement, guarantee, or warranty of any such equipment by the Department of Agriculture.

2/ The author is a Biometrician, In Charge of the Livestock Research Staff of Biometrical Services, ARS, Agricultural Research Center, Beltsville, Maryland.

Data to be processed through ECM equipment must be punched in the card according to a standard arrangement. Consequently, columns on the card are grouped and reserved for the recording of each fact about a record. For example, the first three columns on each card in a particular deck might be reserved for animal or plant number, the next two for year the record was obtained, the next two for month the record was collected,. etc. The set of columns reserved for a particular kind of information is known as a "field."

By giving the specifications desired such as "field" headings, etc., a special plate can be made to produce a card of any type. A charge is made for the cutting of the plate and, therefore, unless the card form is to remain standard and will involve many unit records over a considerable time period, it is usually desirable, in processing research data, to use the standard manilla cards.

The Mark Sense Cards

The more common mark sense card (Figure 2) contains 27 vertical columns and 12 rows. It is the same size as the standard punch card referred to above. The reduction in number of columns is required in order to allow space for placing the marks.

The purpose of the mark sense card is to save time and money in getting records punched onto cards. The mark sense card is first laid off into fields for different types of information in the record and the cards are then marked according to this plan. Although the cards must be marked with a special pencil, the marking may be done at the time the data are obtained. No additional record of the data is necessary. The mark sense cards can be marked from data sheets, if desired, but in that case it is usually more convenient to punch directly from the data sheets into cards. Once the mark sense card has been properly marked the punching is done automatically by machine.

At this point, it should be pointed out that the accuracy of results obtained from processing data with ECM equipment depends very largely on the accuracy of the information punched into the original unit record cards. Before the actual punching of the holes in the cards, it is generally a good practice to double check the accuracy of information recorded on either data sheets (source documents) or on mark sense cards.

Card Punching and Duplicating

The manual punching of cards is the basic method of converting unit records onto punch cards. The operator reads the unit record and, by depressing keys, converts the information into punched holes according to the plan for that particular deck of cards. The machine (Figure 3) feeds, positions, and ejects the card automatically. The operator must always be careful to depress the proper keys in the correct sequence.

This is basically the same kind of function as typing or other key-driven operations. Most operators have little difficulty in punching 100 cards an hour.

Repetitive information is automatically punched from the first card of a group into each succeeding card in that group, as illustrated in Figure 3. For example, if several cards are to be punched for each animal or plant and the data on the source document are arranged by animal or plant number then the identifying information 3/ will be manually punched for the first card for each animal or plant but this information will be duplicated automatically on all other cards. This duplication is normally performed as part of the card-punching function.

The automatic duplication feature on the card punch (see Card Duplicator, Figure 3) is also commonly used to reproduce individual cards that become torn or bent in normal operations of other machines. If cards are properly placed into the hoppers of ECM equipment the bending and tearing of cards seldom occurs. However, when cards do become mangled the automatic duplication feature on the card punch is invaluable since the ordinary reproducer will not process bent or torn cards.

Card Verifying

Card verifying is simply a means of checking the accuracy of the original key punching. A second operator is normally used to verify the original punching, in order to avoid making the same mistakes of reading and to detect errors in punching. The original punching is verified by depressing the keys of a verifier while reading from the same source document. The Machine (Figure 4) compares the key depressed with the hole already punched in the card. A difference causes the machine to stop, indicating a discrepancy between the two operations.

The new verifying machine places a notch in the upper right edge of the card to indicate that it has been key punched and correctly verified. A notch is placed directly above a column that is found to be punched wrong. The card can then be easily corrected.

The accuracy of punching cards is sometimes verified by having a second operator punch a second deck of cards and then comparing the two decks on the collator. When the number of cards is not too great and time on a collator is available, this method of verifying has some advantages over the ordinary verifier.

Interpreting

Some card-punches are equipped with printing mechanisms which automatically

---

3/ Identifying information refers to all of the data necessary to distinguish one card from all other in that deck.

interpret (by printing) the punched information at the top of the card directly above the hole being punched. However, another machine known as the "Interpreter" (Figure 5) is the primary machine used to print the punched information on the top (or bottom) of the card. The interpreter allows the information from any field in the card to be printed in many different positions on that card. Interpreting is advantageous when punched cards are used as documents on which additional information is written or marked, or whenever reference to filing operations is involved.

Reproducing and Gang Punching

Reproducing is the process of automatically punching information from one set of punched cards onto another set of cards. The two sets of cards are fed through the machine (Figure 6) synchronously.

The reproducer allows a considerable amount of flexibility in the reproduction process. Any part or all of the information punched in a deck of cards can be transferred to any section of another deck of cards in any desired sequence. The reproducer is equipped with comparing relays which are used to check the agreement between the originals and reproductions before the cards enter the output hoppers. If a disagreement is found in any one column the machine stops and a comparing light comes on.

The reproducer is also used in gang-punch operations. Gang punching is the automatic copying of punched information from a master card (such as the first card of a group) into one or more detail cards that follow. When a new master card enters the machine the information being gang punched from the previous master card changes to that which is carried in the columns on the new master card being gang punched. In single master-card gang punching, one master card precedes all detail cards to be punched with the same information. The number of columns that can be gang punched is limited only by the number of columns on the card.

Gang punching may be performed separately or in combination with reproducing and summary punching (to be described later) for both alphabetical and numerical information.

Sorting

Sorting is the process of arranging a deck of cards in numerical or alphabetical sequence according to any classification. To arrange cards by animal or plant number, for instance, they are sorted into sequence. Ordinarily, the cards are arranged in ascending rather than descending sequence.

Although the sorting operation is performed on only one column at a time the deck of cards can be arranged into sequence on as many columns as required by a within-within type of sorting. For example, if the cards are

to be arranged into ascending sequence on a three digit field, say columns 1, 2 and 3, then the deck would first be sorted on column 3, then on column 2, and finally on column 1. After each of the first two sorts, the cards would be placed back into the sorter in ascending sequence.

The sorter (Figure 7) provides a fast automatic means of arranging cards for the summarization of data by the tabulating machine. The same deck of cards can be rearranged in any manner desired to obtain different tabulations.

The sorter can be used to select certain cards by allowing it to sort on only one or some combination of the twelve column digits. Sorters are also available which are equipped with counting mechanisms, permitting the counting of digits in a particular column even without any actual sorting of the cards.

Collating

The collator is a machine (Figure 8) that is much more flexible than the sorter, and thus more useful, for selection of certain cards from a deck. For example, the first or last card of a group (such as all cards of the same sire or variety in the same year) can be selected with the collator.

The collator also performs matching and merging processes. An example of a matching operation on the collator is the selection of parent and offspring comparisons from a set of animal data. Two identical decks of detail cards are first prepared. One deck is sorted according to animal (offspring) number and the other deck sorted according to the parent number (either sire or dam). The two decks are then fed into the collator and the animal number of the first deck is matched with the parent number of the second deck and all matched cards are selected. The card for the offspring which has a dam- or sire-record falls into one pocket and the card for the dam or sire falls into another pocket. Unmatched cards of each of the two decks fall into the two other pockets.

Merging is the combining of two sets of punched cards into one set of a given sequence. Both files of cards must first be in the same sequence before they are merged. This function makes possible automatic merging of new cards into an existing deck of cards in sequence on any number of columns with only one pass through the collator. The sequence of the primary deck of cards can be checked simultaneously.

Tabulating

The machine which transforms the punched holes into letters and numbers and prints or summaries from them on paper, is the tabulating or accounting machine (Figure 9). This machine adds or subtracts but will not directly multiply or divide. The modern accounting machines are highly flexible allowing many different kinds of operations to be performed

which involve selection, addition, subtraction, and final printing. Any information that can be printed from the tabulator can also be punched at the same time on summary cards with the aid of a reproducer or summary punch.

The tabulator will add information from several fields for a particular group of cards, print the totals along with identifying information, clear the counter units, and then begin adding the same fields for the next group of cards. If some cards are to be subtracted or not considered at all, the machine can select these cards according to identifying prepunched information.

Two fields on the same card can be added or one subtracted from the other, if desired, and the result immediately printed or retained until a group is completed. This type of operation is called "cross-footing" and can be extended to cover several fields, some being subtracted while others are added.

Information from each card can be printed as the cards pass through the machine. This type of operation is called "listing" and can be combined with the accumulation of totals, selection of cards for printing, etc.

Group printing is the more common operation for the accounting machine. This is the function that summarizes groups of cards and prints the totals. The sequence of the cards as they enter the accounting machine and the operations the machine is wired to perform on those cards determines the type of information obtained. Any number of different types of summaries may be obtained from the same deck of cards by rearranging the cards and rewiring the control panel on the tabulator.

Sums of squares and crossproducts are obtained on the accounting machine by a method of progressive addition. Several modified procedures of this method have been developed for calculating sums of squares and crossproducts on the tabulator. The simplest and most efficient procedure for most statistical problems utilizes "digit" cards and the total transfer feature on the accounting machine. Even then, the tabulator is inefficient for obtaining sums of squares and crossproducts when the number of cards is small.

Calculating

Many types of calculators are manufactured. The primary difference between these machines is the speed with which they can perform the operations of multiplication, division, addition, and subtraction. Operations of this kind that can be done on electrical calculating punches will be described.

These mechanical calculators (Figure 10) are programmed to perform any one or any combination of the arithmetic steps on each card or a group

of cards as they pass through the machine. Factors to be calculated are mechanically read from each card, or series of cards, emitted by a device within the machine, or developed by the accumulation of a series of calculations. One or several results are punched in each card or in a trailer card which follows a group of cards carrying the factors used in the steps of calculation.

Many procedures allow automatic checking to prove the accuracy of calculations. For example, to check the punched result, an AxB calculation can be cross-proofed against a BxA calculation, using different units in the machine, during the same run.

Results from calculations performed on such a machine must be printed and summarized on the tabulator. Most problems involved in the processing of research data utilize the punch, verifier, sorter, collator, tabulator, reproducer, and calculating punch in some manner before completion.

Examples of Use of ECM Equipment

Example 1

An experiment was conducted over a period of several years to determine the influence of supplemental vitamin A on semen quality of young Jersey and Holstein bulls. The cow herds of about 25 Jersey and 25 Holstein cows were divided at random into three groups. The Group I cows received the regular herd ration, Group II received supplemental beta carotene, and Group III received supplemental fish liver oil. Animals born from these cows were retained on the same rations as their dams.

When possible, duplicate semen samples were collected weekly within a thirty minute period from the male progeny, beginning at as early an age as possible and continuing through 16 months of age. In most cases, the following seven semen quality measurements were made on each sample:

(1) Volume in ml
(2) Appearance score
(3) Initial motility score
(4) Concentration of sperm
(5) Percentage of abnormal sperm
(6) Percentage of unstained sperm
(7) Daily motility scores of diluted semen stored at 38° - 42°F.

A total of 42 bulls completed this experiment, of which 23 were Jerseys and 19 were Holsteins. A total of 2,240 samples were collected from the 42 bulls.

Information Desired

(1) Average daily drop in motility for each sample from first to sixth day or until motility ceased.
(2) Linear regression of motility (from first to sixth day of storage) on days stored for each sample.
(3) Obtain natural logarithm for the concentration of sperm for each sample by matching the detail deck of punched cards (one for each of the 2,240 samples) with a natural logarithm deck.
(4) Means of all semen quality measurements by rations, breeds, months, age of bull (months), breed by ration, breed by month of collection, and breed by age of bull.

(5) Fft a parabolic curve to show the change of each of the eight semen quality measurements with age for each of the 42 bulls. Test the significance of the differences between the curves for bulls of the same breed on the same ration for the same quality measurement. Calculate the statistics required to plot the intra-bull curve for each of these 48 breed-ration-quality measurement combinations.

(6) Covariance analysis for each breed separately to test the significance of differences between rations for each semen quality measurement after adjusting for the average influence of age of bull.
(7) Determine the changes in semen quality from month-to-month for each breed separately by adjusting for average effects of age by least squares. Adjust the two measures of sperm livability (obtained in (1) and (2) above) for the influence of first-day motility, as well as age.

Example 2

An interspecific cross was made between the Sioux and Hirsutum varieties of tomatoes. The progeny from both backcrosses were obtained, even though the F_1's could not be crossed to produce an F_2. Progeny from both the Sioux and Hirsutum parents were tested in the same year with the F_1 and backcross progeny. Forty-nine characters were measured on over two thousand plants. For one reason or another, all characters were not measured on all plants and the number of plants in each group, viz, Sioux parents, Sioux progeny, Hirsutum parents, Hirsutum progeny, F_1, backcross to Sioux, and backcross to Hirsutum, varied widely.

Information Desired

(1) Summaries for each character by parent and progeny groups. Number and percentage of progeny in each category will be desired for discontinuous characters such as color, number, etc. Means will be desired by progeny groups for such continuous characters as leaf width, leaf length, seed size, etc. Frequency distributions and standard deviations will also be obtained for each parent and progeny group for all continuous characters.
(2) R X C contingency tables for all two-way comparisons between the discontinuous characters for each parent and each progeny group. These tables are necessary to determine the linkage relationships among these characters.
(3) Correlation coefficients among all comparisons between continuous characters. These tables will be required for each parent and each progeny group.
(4) Prepare tables showing means for all comparisons between continuous characters. These tables will be required for each parent and each progeny group.

Example 3

A factorial experiment was conducted with beef cattle to determine the value of three ratios of grain to hay in combination with two methods of feeding and two levels of protein. Five steers were fed each of the 12

kinds of rations. Initial weights of the steers were obtained and weighings were made at 28-day intervals thereafter until the steers had been fed for a period of 140 days. Type scores of the steers were obtained at the end of the 140-day feeding period. Also slaughter weight, chilled carcass weight, and carcass grade were recorded on each steer. The same experiment was conducted in two years.

Information Desired

(1) An analysis of variance is needed for the gain in weight for each of the five 28-day periods, the average daily gain for the 140-day period, type score, dressing percentage, carcass weight, and carcass grade.
(2) Averages are required for all of these measurements (a) by years, (b) by ratios of grain to hay, (c) by methods of feeding, (d) by level of protein, and by each two-way and three-way subclass of the four classifications.

Electrical-Computing-Machine Procedures

The information desired in this example could be obtained in a number of different ways with ECM equipment. The steps outlined below are general and are given primarily for illustrative purposes for those persons interested in the more detailed operations.

(1) The first step is to set up a card layout for the punching of the detail data. One card will be punched for each steer. Therefore, this problem involves only 120 detail cards (5 steers x 12 rations x 2 years = 120).

Card layout for detail cards:

Information	Columns
Steer Number	1, 2
Year	3 (Coded)
Ratio of grain to hay	4 (coded)
Method of feeding	5 (coded)
Level of protein	6 (coded)
Initial weight	7, 8, 9
First 28-day weight	10, 11, 12,
Second 28-day weight	13, 14, 15,
Third 28-day weight	16, 17, 18,
Fourth 28-day weight	19, 20, 21,
Fifth 28-day weight	22, 23, 24,
Type score	25, 26,
Slaughter weight	27, 28, 29,
Chilled carcass weight	30, 31, 32,
Carcass grade	33, 34

(2) Punch and verify the detail cards according to the above card layout.

(3) Calculate average daily gain for each steer by subtracting initial weight (cols. 7, 8, 9) from the weight at the end of the fith 28-day period (cols. 22, 23, 24) and dividing the result by 140. Punch the result in columns 35, 36, 37, carrying two decimals.

(4) Calculate the dressing percentage by dividing the chilled carcass weight (cols. 30, 31, 32) by slaughter weight (cols. 27, 28, 29) and multiplying by 100. Round off to one decimal and punch the result in columns 38, 39, 40.

(5) Sort the 120 detail cards into ascending sequence by years.

(6) Tabulate and summary punch totals for card count, each 28-day weight, type score, carcass weight, carcass grade, average daily gain, and dressing percentage along with identifying information. Gang punch the sort number in columns 41 and 42.

(7) Repeat (5) and (6) successively--sorting by ratios, method of feeding, level of protein, years x ratio, years x method of feeding, years x level of protein, method of feeding x level of protein, and by each of the four three-way combinations of these four classifications.

(8) Prepare one total summary card that will contain the totals of all fields for all 120 cards.

(9) A total of 84 summary cards will now be punched from 15 different sorts. Reproduce these into the card form shown below so that one card will be produced for each total on each summary card. Since there will be 10 totals on each of the 84 summary cards the total number of working cards reproduced from these into the card form shown will be 10 x 84 or 840.

Card Layout for working deck of summary cards:

Information	Columns
Variable number	1 (coded and gang punched)
Sort number	2, 3
Card count (number of steers in this classification or subclass)	4, 5, 6,
Total (Y)	7, -12,

(10) Calculate the mean (total (Y) divided by card count) and the correction factor (total (Y) squared divided by card count), and in the same operation punch these results on each of the 840 cards.

(11) List and summarize totals of card count, the Y field, and the correction-term field on the tabulator by variable and sort number. In the same operation, subtract the total correction term for the respective variable from the sum of the individual correction terms within each sort and print out the result. These quantities will be the corrected "between" sum of squares for each classification and for each of the two- and three-way subclasses.

(12) Calculate the total sum of squares for each of the 10 variables on the 602-A Calculating Punch, punching the result in each case on a trailer card. This operation will require 10 passes through the machine of the 120 detail cards. An additional 10 passes will be required for checking the computations.

All the sums of squares necessary to set up the 10 analysis-of-variance tables will now be available and little hand computation will be required to complete the analyses. All 840 means will appear on the tabulation.

A research problem of this magnitude is frequently done with a desk calculator. However, where a routine ECM procedure has been established with permanently-wired control panels allowing considerable flexibility in the number of variables, size of fields, etc., then the problem can be efficiently handled with a saving of time and expense to the investigator.

The procedure outline above, to handle this analysis of variance problem, needs only slight modification in order to be used equally well for correlation, regression, and covariance problems.

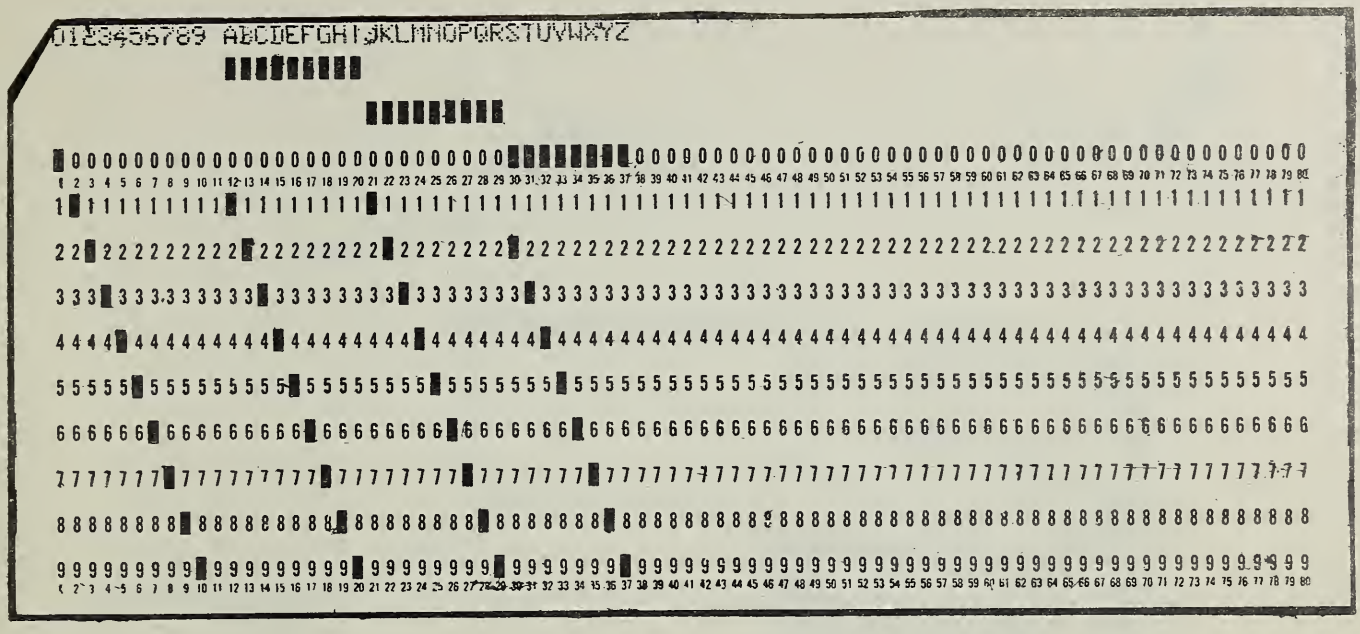


Figure 1 STANDARD PUNCH CARD

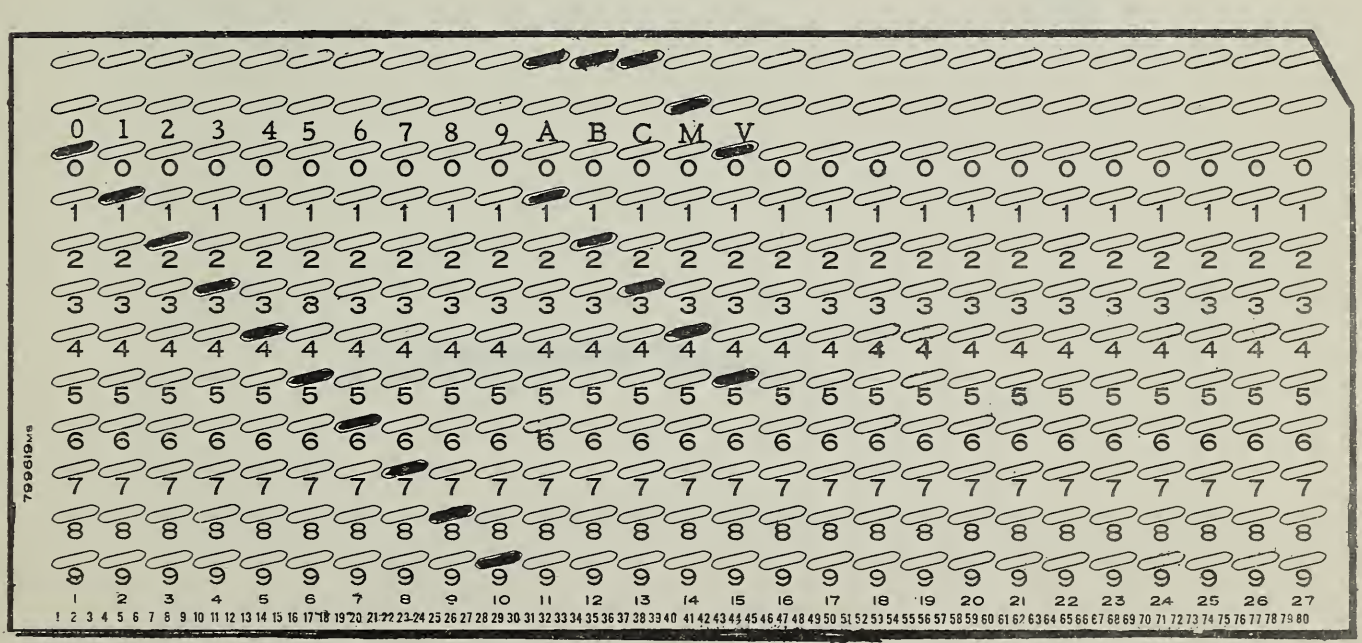
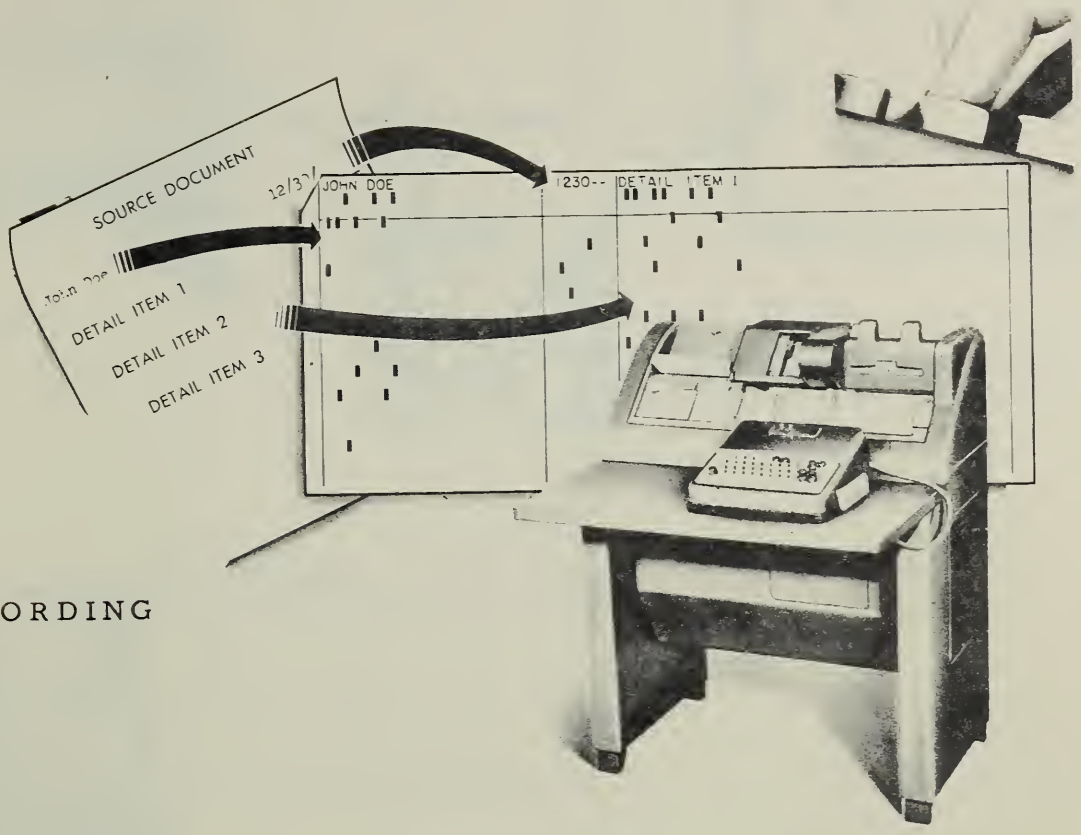
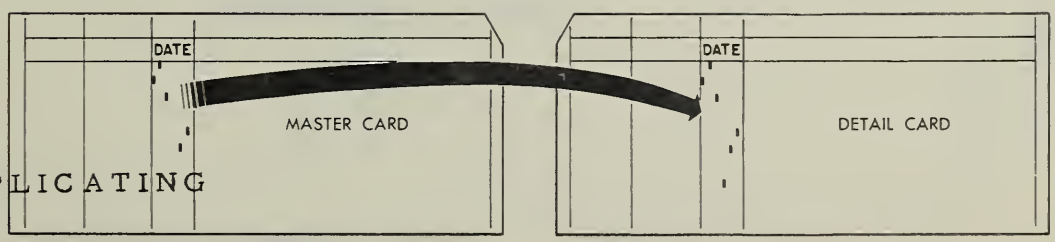


Figure 2 MARK SENSE CARD



RECORDING



DUPLICATING

Figure 3 CARD PUNCH FOR RECORDING DATA IN A CARD AND FOR CARD DUPLICATION.

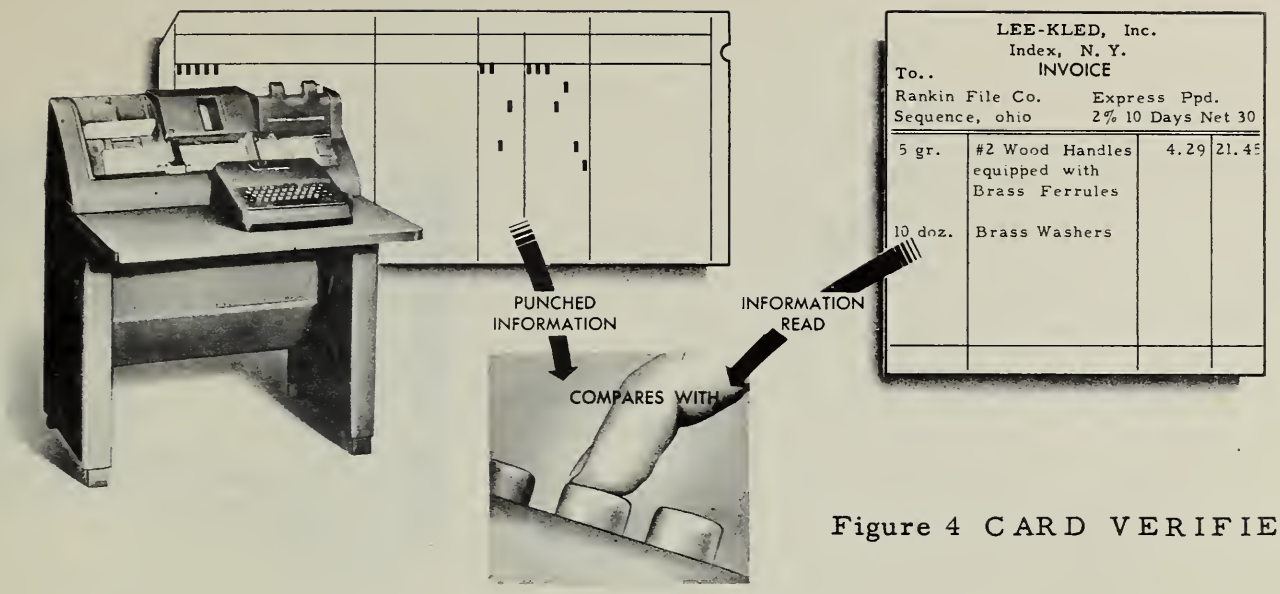


Figure 4 CARD VERIFIER

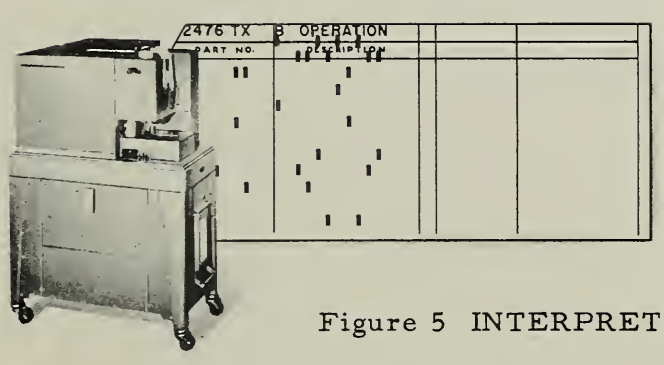


Figure 5 INTERPRETER

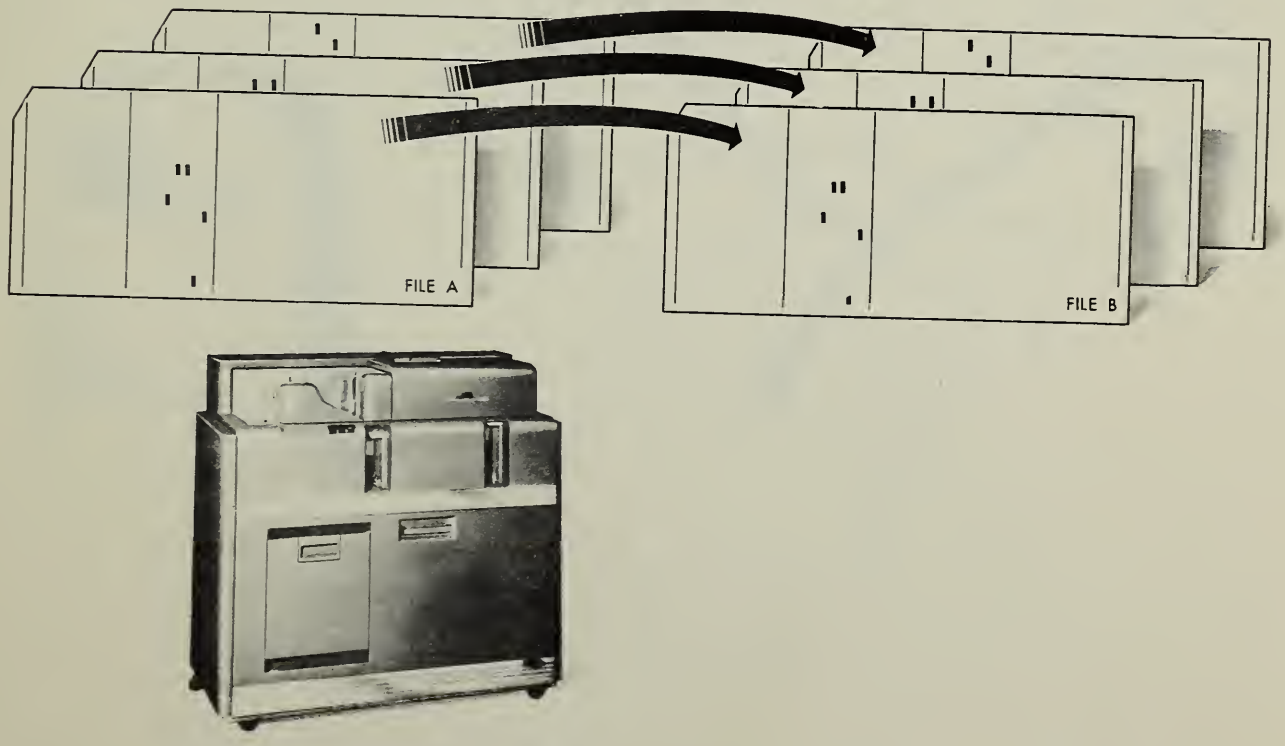
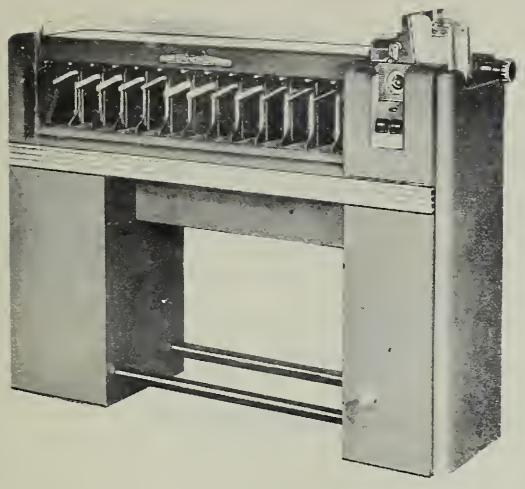


Figure 6 REPRODUCER



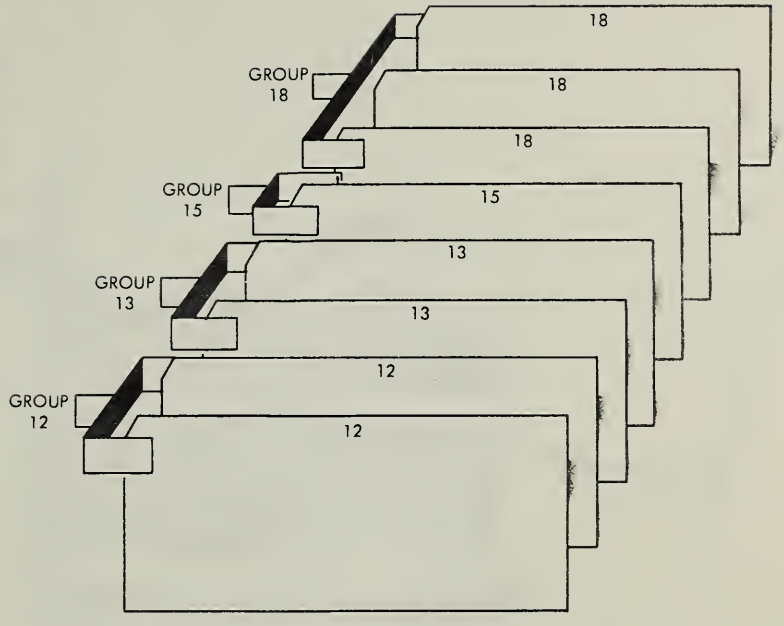


Figure 7 SORTER

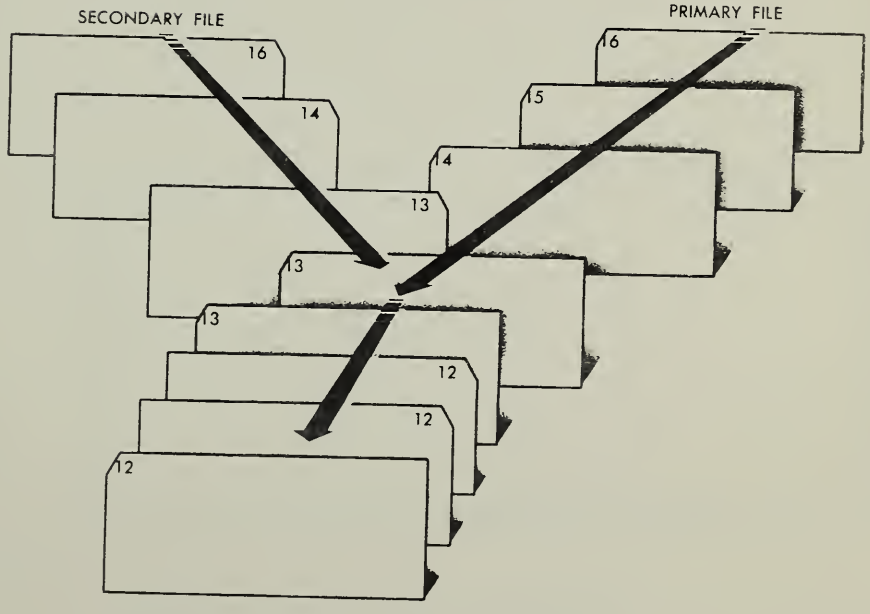


Figure 8 COLLATOR

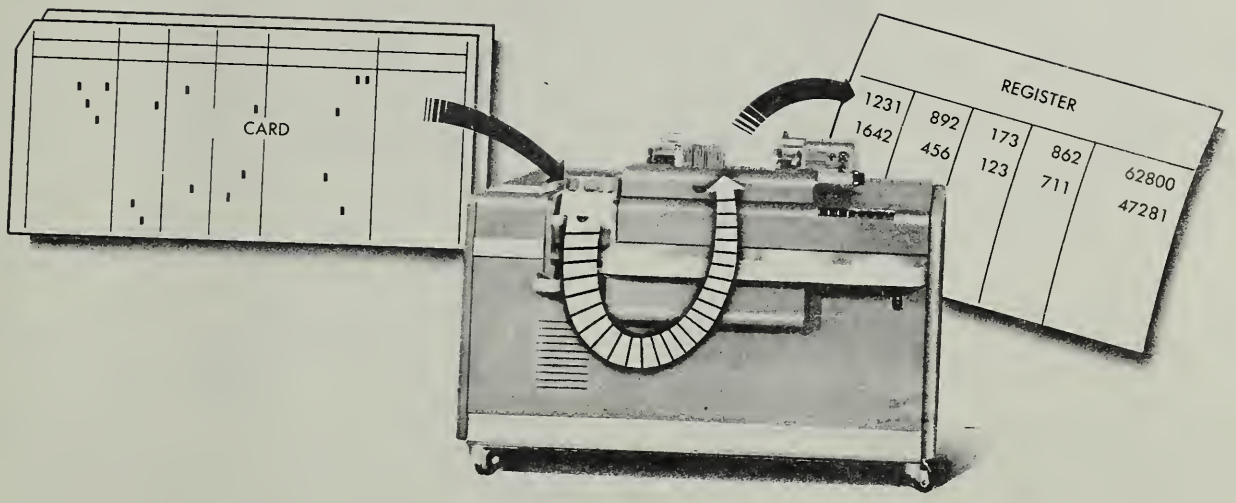


Figure 9 TABULATOR

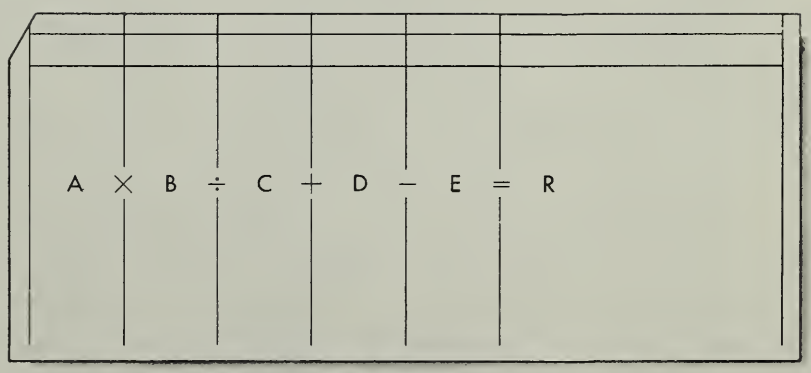


Figure 10 CALCULATOR